Crystal Structure Predictions for Acetic Acid

WIJNAND T. M. MOOIJ,[1] BOUKE P. VAN EIJCK,[1] SARAH L. PRICE,[2] PAUL VERWER,[3] JAN KROON[1]

[1] *Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*
[2] *Chemistry Department, University College London, Christopher Ingold Laboratories, London, UK*
[3] *CAOS/CAMM Center, University of Nijmegen, Nijmegen, The Netherlands*

Received 31 March 1997; accepted 8 October 1997

ABSTRACT: Possible crystal structures of acetic acid were generated, considering eight space groups and assuming one molecule in the asymmetric unit. Our grid-search method was compared with a Monte Carlo approach as implemented in the Biosym/MSI Polymorph Predictor. This revealed no sampling deficiencies. A large number of possible crystal structures were found (~ 100 within only 5 kJ/mol), including the experimental structure. Energy minimizations were done with a united-atoms force field (GROMOS), an all-atoms force field (AMBER), and a potential that describes the electrostatic interactions with distributed multipoles (DMA). In all cases, the experimental structure had a low lattice energy. The number of hypothetical crystal structures was reduced considerably by removing space-group symmetry constraints, or by a primitive molecular dynamics shake-up. Nevertheless, sufficient structures of equal or lower energy compared with the experimental structure remained to suggest that other factors need to be considered for genuine structure prediction.
 J Comput Chem 19: 459–474, 1998

Keywords: crystal structure prediction; distributed multipoles; molecular dynamics; symmetry constraints

Correspondence to: W. T. M. Mooij; e-mail: w.t.m.mooij@chem.ruu.nl

Contract/grant sponsors: The Netherlands Foundation for Chemical Research; The Netherlands Organization for Scientific Research; contract/grant number: 96PPM 012

This article contains Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/19/459 or http://journals.wiley.com/jcc/

 CCC 0192-8651 / 98 / 040459-16

Introduction

Over the past few years there has been an increasing interest in the prediction of crystal structures on the basis of molecular information only. Different methods have been developed to generate possible crystal structures,[1–17] an overview of which can be found in ref. 18. All studies agree that there will often be a large number of possible crystal structures, as judged from their lattice energy, and nearly all are successful in finding the experimental structure. The experimentally observed structure is generally very close in energy to the global minimum in the lattice energy.

We have previously used a brute-force grid-search method (in the space group $P2_12_12_1$) for the prediction of the crystal structures of six monosaccharides.[6] In this study we test the ability of our method to predict correctly the crystal structure of acetic acid.[19] This structure is an interesting one, as it shows a catemer motif as opposed to almost all other carboxylic acids, which crystallize as cyclic dimers. Almost 20 years ago, Derissen and Smit[20] searched for possible dimer structures (using fixed cell dimensions in space group $P2_1/c$), and found two candidates with a lattice energy about 1 kJ/mol higher than the experimental structure. Now a systematic *ab initio* search is feasible. Our method uses a systematic search of the parameter space, whereas many other published methods rely on some kind of random search. To compare the two strategies, we also employed the commercially available Biosym/Molecular Simulations Polymorph Predictor[2–4,21–23] to generate possible crystal structures for acetic acid.

The choice of the force field may be more important than the search strategy. In the absence of comparative studies it is of interest to see whether different force fields will produce the same predictions for the same compound. Based on the result of the grid search, a united-atom (GROMOS),[24] all-atom (AMBER),[25] and a much more sophisticated, DMA-based force field are used. The DMA-based force field uses distributed multipoles, calculated from an *ab initio* wave function, to give an accurate description of the electrostatic interactions.

We also investigate the effect of removing the space-group symmetry constraints (except for the existence of a periodic lattice) on the predictions, and the use of molecular dynamics to reduce the number of candidate crystal structures.

Method

UPACK SEARCH AND ANALYSIS

The program,[6] which performs a brute-force grid search for possible crystal structures, has been completely revised and will be referred to as Utrecht crystal PACKer (UPACK). The code has been altered to handle triclinic, monoclinic, and orthorhombic space groups to be able to search the most abundant space groups for organic molecules. It is still limited to one molecule in the asymmetric unit. The crystallographic a-axis is chosen to coincide with the Cartesian x-axis, and the b-axis lies in the $x-y$ plane. Then the crystallographic parameters to be varied are the orthogonal components of the cell axes (a, b_y, and c_z) and the components b_x, c_x and c_y, which determine the cell angles. For a monoclinic structure, b_x and c_y are just zero; for an orthorhombic structure, c_x also is zero. The search is limited to densely packed structures by calculating each starting value of c_z from an estimated cell volume ($c_z = V_{\mathrm{est}}/ab_y$), which is not very critical. The component b_x has to be varied from 0 up to the length of a to find all unique settings of the $a-b$ plane. Likewise, c_x has to be varied from 0 to a and c_y from 0 to b_y. Into each trial cell one independent rigid molecule is placed, with center-of-mass coordinates X, Y, Z (to be varied if necessary in the space group under consideration) and varying Euler angles (φ, ϑ, ψ).

The search proceeds by optimizing these starting parameters roughly, using a few simple steepest descent steps followed by some regular conjugate gradient minimization.[26] Equivalent structures must be removed from these lists. Because these may differ in choice of unit cell, it is not always immediately obvious whether or not two structures are equivalent, a problem for which we recently developed our own fast clustering algorithm.[27] Briefly, this algorithm tests for equivalence between two structures by checking whether or not the rotation necessary to transform one orientation into the other corresponds to a cell transformation allowed by the space-group symmetry. This test allows for deviations of magnitude ϵ_N from integer values in the transformation matrix elements. A final test compares all the com-

ponents of the cell axes and the molecular centers of gravity after the transformation of one structure into the other, allowing for a tolerance, ϵ_A.

The second stage of the procedure consists of a further rigid-body refinement. To reduce the computational effort it is worthwhile to perform an intermediate round of clustering when the structures are not yet properly minimized. This is followed by the final minimization, until a proper lattice-energy minimum is reached. The resulting lists of possible crystal structures are once more clustered. This is extremely important to limit the number of possible crystal structures, but clustering will be hindered by structures that have not yet completely converged to the same minimum. So, the final energy minimization must be continued until complete convergence is reached, for which it is necessary to use double precision and a very sensitive convergence criterion (0.001 kJ/Å or less for the rms gradient of the energy). As a cut-off distance for summing the intermolecular interactions is used, a fixed pair list is necessary to avoid discontinuities in the energy between successive minimization steps. This pair list is updated only when a parameter has changed more than a certain threshold. The rigid-body minimization can be followed by a minimization allowing for full geometry relaxation.

For the polar space groups, $P2_1$ and $Pna2_1$ in this work, there are additional problems in calculating the lattice energy,[28] which are now handled in UPACK. Briefly, if the energy of a certain volume, V, with a dipole moment, p, is calculated in a straightforward manner using a cut-off radius, the term $2\pi p^2/3V$ should be subtracted to obtain the same result as from Ewald summation without the $h = 0$ term.[29] In both approaches, an additional energy contribution that depends on the external crystal form is needed, which may be neglected if the charges on the crystal surface are counterbalanced by some external charges or if the crystal shape corresponds to one with the lowest possible energy. This is assumed throughout this study.

Structures that are found after the final minimization are minima in the energy surface under the constraints of the imposed space-group symmetry. They are not necessarily minima in the absence of those constraints; that is, the crystal does not have to be at minimum energy with respect to the extra degrees of freedom. When the space-group constraints are released by expanding the structures to $P1$ all the forces on the molecules remain zero by reasons of symmetry, including the forces in the direction of the extra degrees of freedom: the structures are minima or saddle points in the complete crystal-energy surface. Whether a structure is a true minimum can be tested by minimizing the crystal structures after an expansion to $P1$. However, starting with perfect symmetry, all derivatives are zero and nothing would happen. Thus, some small random shifts (e.g., 0.05 Å) are needed to destroy the perfectness of the initial symmetry.

After minimizing, the structures must be analyzed for the remaining symmetry using a program like PLATON.[30] Clustering of these structures, which may have several independent molecules ($Z' > 1$), cannot be performed with the clustering algorithm.[27] An atom-atom distance-based clustering strategy has been adopted for these cases. For each molecule in the $P1$ unit cell a list of atom−atom distances within a certain distance (ρ_c) is made. First, within one crystal structure the lists for all the separate molecules are compared; all distances need to be identical within a given tolerance (η) for the molecules to be considered symmetry related. This gives the number of independent molecules in the unit cell, albeit based on local rather than crystallographic symmetry. Second, different structures which have been assigned the same number of independent molecules are compared, again allowing for a maximum difference η in the distances. This code is considerably slower than the clustering algorithm,[27] but this is not a problem as the number of structures to be considered at this stage is not excessively large.

It is conceivable that two energy minima are separated by a fairly low barrier. In such a case the least favorable one can easily convert into the other, thus reducing the number of possible crystal structures. To investigate this possibility we subject the structures to a primitive molecular dynamics simulation, where the space-group symmetry is enforced at every time step. This is, of course, very unrealistic, but any other approach would make excessive demands on computer time. These simulations are just meant to ''shake up'' the structures to overcome barriers of the order of kT. Simulations were done at constant temperature (200 K, but starting from 10 K) and pressure (1 atm), using the weak-coupling method.[24]

POLYMORPH PREDICTOR SEARCH

A crystal-packing prediction with the Polymorph Predictor consists of a sequence of four

steps that can be carried out automatically. It produces a set of predicted crystal structures for a single space group with a given number of independent molecules in the unit cell, and a given starting molecular geometry. The program can handle all 230 space groups, with any number of independent molecules in the cell.

In the first stage of the prediction, a Monte Carlo/simulated annealing packing algorithm generates starting structures, treating the molecule as a rigid unit. In each Monte Carlo step, all relevant parameters are changed randomly. To suppress crystal evaporation the parameters defining the spatial extension of the crystal are handled in a different manner, as specified in the literature.[2] A new structure is then accepted or rejected using the standard Metropolis test. The Monte Carlo simulation takes 4000–5000 steps, resulting in about 2000 accepted structures. The simulated annealing procedure is applied to improve the efficiency of finding the most promising starting structures. During the Monte Carlo simulation, the temperature parameter is first increased after each step, until a specified number of consecutive trials (usually about 10) have been accepted. This will take 50–100 steps, and the temperature parameter will rise to about 10,000 K. During the rest of the Monte Carlo simulation, the temperature is slowly decreased until it reaches a certain minimum value (usually 200 K).

The second stage of the prediction consists of clustering of the structures produced by the Monte Carlo run. Clustering is based on interatomic distances, discriminating by force field type. Because six different atom types are present in the acetic acid molecule (using the DREIDING[31] force field), 21 different combinations of two atom types are possible, and thus 21 types of interatomic distances are present. For each structure a list of interatomic distances within a certain cut-off is made for each combination of atom types. By comparing these lists, the dissimilarity between structures is calculated, which is used to decide whether structures belong to the same cluster or not. The number of clusters formed in this case is around 1000.

The third stage of the calculations consists of a full geometry minimization of one structure from each cluster, optimizing atomic coordinates and cell parameters, but imposing space-group symmetry. Finally, a clustering is performed on all minimized structures, identical to the clustering just described, leaving somewhere between 10 and 300 different structures, within an energy range of about 40 kJ/mol.

FORCE FIELDS

GROMOS Force Field

The standard GROMOS force field[24] was used with charges +0.15 e for the united CH_3 "atom," +0.38 e for C, −0.38 e for the carbonyl O, −0.55 e for the hydroxyl O, and +0.40 e for the hydroxyl H.

AMBER Force Field

As an example of an all-atom force field, AMBER[25] was used with charges −0.22 e for the CH_3 carbon, +0.10 e for the CH_3 hydrogens, +0.55 e for the carbonyl C, −0.50 e for the carbonyl O, −0.58 e for the hydroxyl O, and +0.45 e for the hydroxyl H. These charges are obtained by adaptation of (united-atom) AMBER/OPLS charges.[32] Two missing angle parameter sets had to be defined. The O=C—O and the C—C—O angle parameters were taken to be 124°, 80 kcal mol^{-1} rad^{-2} and 115°, 70 kcal mol^{-1} rad^{-2}, respectively, similar to both the GROMOS values and the COO^- parameters in AMBER.

DMA-Based Force Field

This potential describes the electrostatic contribution to the lattice energy very accurately as the charge distribution is not just described by partial atomic charges, but by sets of atomic multipoles. The electrostatic interactions are evaluated using all terms up to R^{-5} in the multipole expansion between the multipoles that are located on all the atoms. Multipoles up to hexadecapole are used on all atoms except hydrogen, where only the charge and dipole are used. The multipoles were obtained from a distributed multipole analysis[33] of the wave function of an isolated acetic acid molecule. The wave function was calculated at the MP2 level using a 6-31G**[34] basis set in the program CADPAC.[35] The wave function is calculated at the geometry used in the predictions; that is, either the experimental geometry or the geometry of the free molecule minimized in the GROMOS force field has been used.

Although the electrostatic model is theoretically well founded, the repulsion and dispersion terms are necessarily modeled by an empirical atom–atom 6-exp potential. The parameter set FIT[36] was

used, as it gives a reasonably good reproduction of the crystal structures of a wide range of polar molecules, including N—H ⋯ O hydrogen-bonded structures, when used in conjunction with a DMA model, and it also gives a good reproduction of the crystal structure of acetic acid. However, this involved transferring the parameters for amide hydrogen to hydroxyl hydrogen atoms, which resulted in somewhat elongated hydrogen bonds (see Table II). In the course of this research it was found empirically that scaling the polar hydrogen pre-exponential factor by 0.85 (justified by the change from N—H to O—H) gave an even better reproduction of the structure, including the hydrogen-bond length. A test set of crystal structure predictions (in the space groups $P\bar{1}$, $Pna2_1$, and $P2_1/c$) was carried out using this potential, denoted DMA(OH). Although this changed the relative energetic order of the crystal structures under consideration, it did not change the nature of the predictions, and thus it was not worth repeating the entire procedure. A more accurate development of the entire dispersion/repulsion potential for carboxylic acids would provide a more definitive potential.

The calculations using the DMA potential were done using the program DMAREL.[37] This program minimizes the lattice energy of a molecular crystal, assuming the molecules to be rigid. Charge–charge, charge–dipole, and dipole–dipole interactions were calculated using the Ewald summation technique. All higher multipole interactions and dispersion/repulsion interactions were calculated using a cut-off radius of 15 Å. As the program cannot constrain space-group symmetry, the minimizations were performed on $P1$ unit cells.

DREIDING Force Field

In the Polymorph Predictor runs, the DREIDING-2.21 force field,[31] as implemented in Cerius2 (Cerius2 is a trademark of Biosym/Molecular Simulations), was used. Calculations were carried out using several types of atomic charges: MNDO[38]/ESP charges, scaled by 1.4, as suggested by Besler et al.,[39] as well as ESP charges derived from *ab initio* calculations using an STO-3G basis set, scaled by an empirical factor. The effect of using different sets of atomic charges or completely different scaling factors on geometry and ranking turned out to be small; the results obtained using STO-3G/ESP charges scaled by 0.7 are given here. The corresponding charges were: CH_3 carbon, −0.3344; CH_3 H1, +0.0797; CH_3 H2, H3, +0.0926; carbonyl C, +0.5340; carbonyl O, −0.3103; hydroxyl O, −0.3793; hydroxyl H, +0.2250 (CH_3 H1 is the hydrogen that is eclipsed with respect to the carbonyl O).

Results

SEARCH RESULTS AND RANKING

UPACK

The search for possible crystal structures was carried out in the eight most abundant space groups, up to the space group of the experimental structure ($Pna2_1$). Together, they cover 85% of the crystal structures in the Cambridge Crystallographic Database. The GROMOS force field was used and the molecular geometry was taken to be the minimized structure of a free acetic acid molecule in that force field. In the grid search the cell parameters and center-of-mass coordinates were varied in steps of 1 Å, and the Euler angles in steps of 30°, using a relatively short cut-off radius of 8 Å to evaluate the lattice energy. The number of initial structures varied from about 16,000 in $P2_1$ to 22,000,000 in $C2/c$.

The numbers of resulting structures ar given as N_0 in Table I. The clustering procedure reduced those numbers to N_1. The tolerances here were $\epsilon_N = 0.2$ and $\epsilon_A = 1$ Å.[27] Throughout the rest of the procedure clustering tolerances of $\epsilon_N = 0.3$ and $\epsilon_A = 0.3$ Å were used.

Continuing the energy minimization with the GROMOS force field the cut-off radius was extended to 10 Å. The numbers of resulting structures after clustering are given as N_2 in Table I. The final rigid-body minimization to a proper energy minimum was done using a cut-off radius of 12 Å. The numbers of structures that remained after clustering are reported as N_3 in Table I. It can be seen that continuing the minimization of the approximately minimized structures (N_2) until complete convergence had been reached (N_3) proved very valuable. This is why the numbers reported here are lower than those reported previously.[27] Acetic acid has only few internal degrees of freedom, especially in the united-atom approximation of GROMOS, so it was surprising to see that the number of structures decreased upon going to a flexible model; that is, after a full geometry relaxation and clustering, the numbers N_4 in Table I remained. Inspection of some structures

TABLE I.
Search Results.[a]

		$P\bar{1}$	$P2_1$	$P2_1/c$	$C2/c$	$P2_12_12_1$	$Pna2_1$	$Pbca$	$Pnma$
N_0		37,705	2294	60,683	69,908	7035	3874	9410	15,100
N_1	(0.2, 1.0)	5225	432	15,051	23,122	872	1111	1838	3971
N_2	(0.2, 1.0)	150	69	1153	1709	162	202	210	338
N_3	(0.3, 0.3)	7	13	137	188	21	58	55	45
N_4	(0.3, 0.3)	7	10	113	139	19	47	49	40
N_5	(0.3, 0.3)	15	15	197	320	25	61	68	67
N_6	(0.3, 0.3)	9	8	117	157	17	41	37	38
N_7		30	20	152	270	32	84	82	107

N_0 = number of structures delivered by the search procedure; N1 = number of structures after clustering; N2 = number of structures after intermediate minimization and clustering; N3 = number of structures after complete rigid-body minimization and clustering in the GROMOS force field; N4 = number of structures after full geometry relaxation and clustering in the GROMOS force field; N5 = number of structures after complete rigid-body minimization and clustering in the AMBER force field; N6 = number of structures after full geometry relaxation and clustering in the AMBER force field; N7 = number of structures resulting from Cerius2 Polymorph Predictor run.

[a] Values in parentheses denote clustering tolerances ϵ_N and ϵ_A, respectively.[27]

that are unstable in a ''flexible'' minimization revealed that they have the acetic acid molecule located on a crystallographic mirror plane. In a rigid-body approximation there apparently existed an energy barrier for rotating the molecule out of this plane, which was absent in the case of a flexible molecule. Indeed, in the first case, a moderate distortion of the symmetry was seen to return to the starting position, showing it to be a true minimum; in the second case, even a very small distortion resulted in a reorientation of the molecule, showing that the structure now represented a saddle point. On visual inspection it was far from obvious how the flexibility was used to remove such an energy barrier.

The prediction using the AMBER force field started from the results of the first stage of the packing procedure. Although these structures were obtained after some minimization in the GROMOS force field, their number was still so large (N_1 in Table I) that we consider it highly unlikely that some potential AMBER structures were missed in the end. The molecular geometry was changed to the minimized structure of a free acetic acid molecule in that force field. The final rigid-body energy minimization was done with a cut-off radius of 12 Å, resulting after clustering in the numbers N_5 in Table I. In this all-atom force field going to a flexible molecule introduces extra degrees of freedom, like a rotatable methyl group. Indeed, the full geometry relaxation (cut-off 12 Å) resulted in a larger decrease in the remaining structures after clustering (N_6 in Table I) than in the united-atom GROMOS force field.

The DMA potential was not used in the search and preliminary minimization stage of the procedure, as that would be computationally too demanding. So, calculations using this potential were done starting from the final GROMOS-based predictions (N_3 in Table I). When this number did not exceed 20 entries for one space group, all those structures were minimized in the DMA-based force field; the longer lists were truncated, using all structures within 15 kJ/mol of the global minimum in the GROMOS force field.

The grid-search algorithm has succeeded in finding the minimum corresponding to the experimental structure and, as a result, the experimental structure was present in the final structure lists of all force fields. However, the structural parameters for the energy-minimized experimental structures were different for the different force fields. The deviations from the true experimental structure were rather large for both the GROMOS and the AMBER potential; the DMA and the DMA(OH) potential had minima that deviated only slightly from the experimental structure. The cell axes are given in Table II, together with the ranking of the experimental structure. It is seen that all potentials resulted in a reasonably low ranking with an energy difference that is of the same order of magnitude as what one could reasonably expect for the uncertainties of the potentials. An overview of the distribution of structures within 10 kJ/mol of

TABLE II.
Cell Axes, Hydrogen-Bond Geometries, and Rankings of Lattice-Energy Minima Corresponding to the Experimental Structure, for Different Force Fields.

	a (Å)	*b* (Å)	*c* (Å)	O ··· O (Å)	O — H ··· O (°)	Ranking[c]	ΔE[d]
Exp.	13.22	3.96	5.76	2.63	165	—	—
GROMOS	14.27	3.94	4.74	2.62	178	11	0.9
AMBER	14.29	3.87	5.30	2.66	171	4	2.1
AMBER[a]	14.25	3.83	5.26	2.63	173	3	0.9
DMA[b]	13.49	3.93	6.03	2.81	162	11	1.3
DMA[a]	13.36	3.90	5.99	2.81	168	7	1.0
DMA(OH)[b]	13.45	3.90	5.86	2.64	158	17[e]	2.8[e]
DMA(OH)[a]	13.24	3.88	5.86	2.63	163	—	—
DREIDING	14.45	4.07	5.48	2.96	170	11	1.0

[a] Experimental molecular geometry used.
[b] GROMOS molecular geometry used.
[c] A ranking of *r* implies that the structure that corresponds to this minimum was the *r*th lowest in energy.
[d] Difference in lattice energy from the global minimum found for all space groups for the given force field.
[e] Result of a very limited search. Starting points were minima in the GROMOS force field in $P\bar{1}$ (7 structures), *Pna*21 (40 structures), and $P2_1/c$ (60 structures).

the global minimum in the lattice energy is given in Figure 1a–c. The AMBER potential gave the fewest structures with an energy close to its global minimum, but differences between these histograms should not be given too much significance: the accuracy of the potentials is not expected to be within 1 kJ/mol.

The lowest lattice energy per space group is given in Table III, together with the number of structures within 5 kJ/mol. The large difference in lattice energy between the DMA potential and the other two force fields came from the hydrogen-bond distance being too large (Table II). The lattice energy of the minimum corresponding to the experimental structure in the DMA(OH) potential was −66.3 kJ/mol. The experimental heat of vaporization (into monomeric gas) is 67 ± 1 kJ/mol at 223 K,[40] from which a lattice energy of −71 kJ/mol was found with the usual 2RT correction. Given the uncertainties in comparing calculated lattice energies with observed sublimation enthalpies,[41] GROMOS, AMBER, and DMA(OH) all gave reasonable values.

The two hypothetical (dimeric) crystal structures of Derissen and Smit[20] were also found. However, they showed some rearrangement upon minimizing in both the AMBER and GROMOS force fields. These two dimeric structures are, however, not particularly highly ranked; furthermore, better dimeric structures were found.

In many crystal structure predictions, rigid molecular geometries are taken from the experimental structure. In a true prediction, of course, this is not allowed. To test whether predictions would change upon this ''cheating,'' the energy-minimized molecular geometry was replaced with the experimental geometry, both within the AMBER and the DMA results. The influence on the minimized experimental crystal structure was marginal (Table II), and only small shifts in rankings of hypothetical structures were observed. One can conclude that the predictions were not very sensitive to the precise rigid molecular geometry used.

Polymorph Predictor

Crystal-structure predictions were carried out in the same eight space groups that were considered in the UPACK calculations. The DREIDING force field, as described previously, was used. So the search strategy as well as the force field were different, allowing a comparison with the UPACK procedure on both levels. The isolated molecule was optimized at the STO-3G level using GAUSSIAN-94.[42] A number of standard settings for all simulation steps in the Polymorph Predictor is available. The Monte Carlo search level was set to ''fine,'' the clustering search level was also set to ''fine,'' the tolerance increased to 0.18, and the limit on the number of generated clusters was removed. All structures were fully minimized, taking the rms force of 0.001 kcal/Å as convergence criterion. van der Waals and Coulomb interactions were calculated using the Ewald summation method.

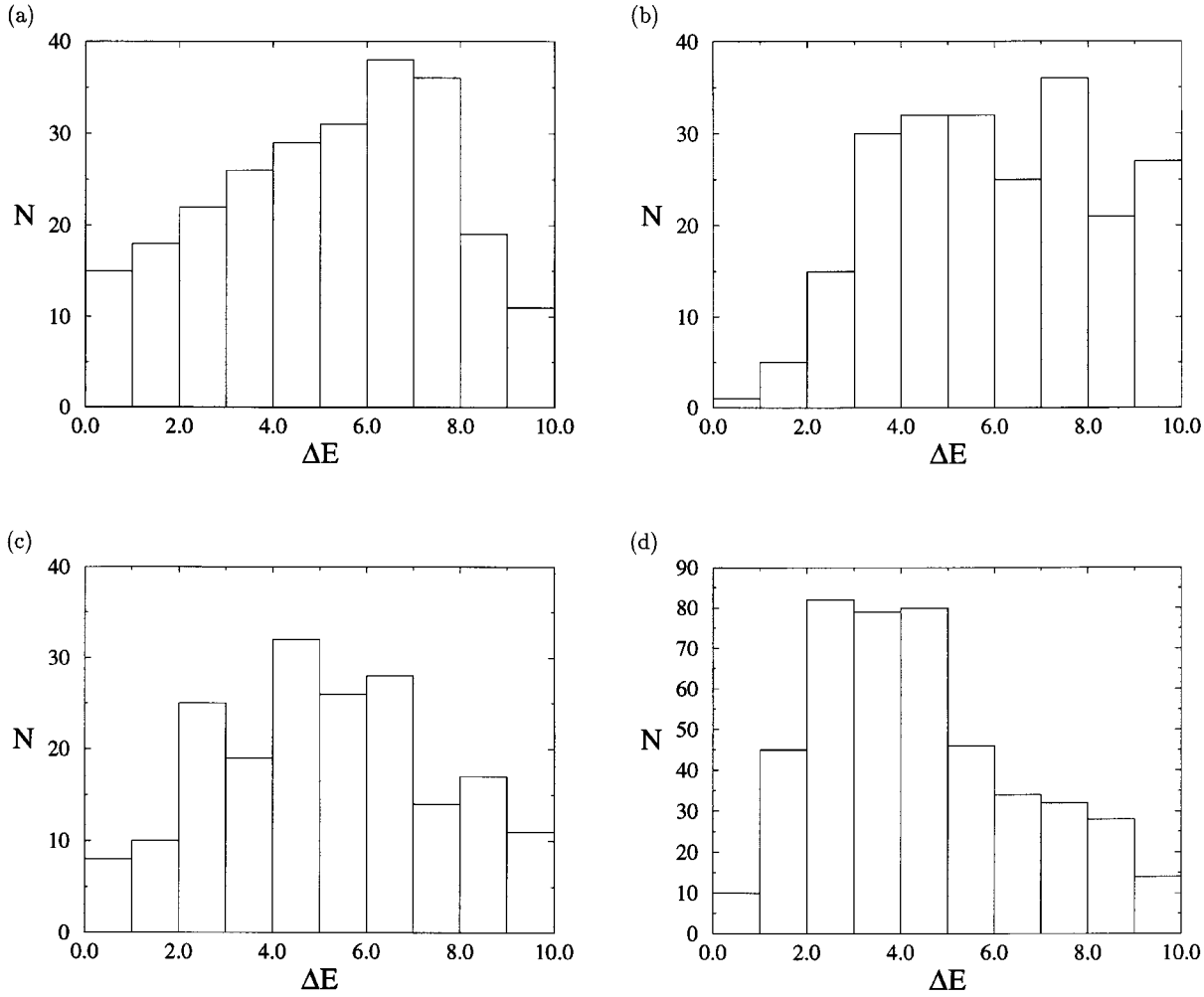

FIGURE 1. Distribution of predicted structures for the different force fields (ΔE is the energy difference with the global minimum (kJ/mol), N is the number of structures). (a) GROMOS force field. (b) AMBER force field. (c) DMA force field. (d) Polymorph Predictor results (DREIDING force field).

TABLE III.
Lattice Energy and Number of Low-Energy Structures.

	$P\bar{1}$	$P2_1$	$P2_1/c$	$C2/c$	$P2_12_12_1$	$Pna2_1$	$Pbca$	$Pnma$
GROMOS[a]	−69.2	−67.3	−69.7	−69.9	−69.8	−69.5	−69.1	−63.5
N	3	4	46	28	11	12	11	0
AMBER[a]	−71.0	−72.6	−75.0	−72.5	−75.0	−75.0	−73.4	−69.4
N	6	3	33	22	7	11	7	0
DMA[a]	−55.3	−53.3	−55.8	−54.9	−55.0	−55.0	−54.2	−49.7
N[b]	4	3	38	35	7	4	9	0
DREIDING[a]	−42.4	−42.2	−43.3	−42.2	−43.3	−43.3	−42.6	−38.8
N	12	10	84	100	21	30	41	6

N = Number of structures within 5 kJ/mol from the first ranked structure for all space groups.
[a] Lattice energy of first ranked structure in that space group. The lattice energy has been calculated as the (intra- plus intermolecular) energy of the molecule in the crystal minus the energy of a minimized free acetic acid molecule.
[b] Only structures that kept their original space-group symmetry.

The numbers of structures that remained after minimization and the final round of clustering are given as N_7 in Table I. Due to a bug in the code, incorrect symmetry elements were generated for a small fraction of the minimized structures in $P2_1/c$ and $C2/c$. After removal of those structures, the energy range for the predicted structures was about 40 kJ/mol.

The experimental structure was found, but the predicted unit cell was significantly larger: axes differed by up to almost 10% (Table II). Another striking difference was the hydrogen-bond distance that was 0.33 Å too large (Table II), which would explain why the structures were too weakly bound (Table III). The energy of the experimental structure was again slightly (1.0 kJ/mol) higher than the global minimum. This most stable structure had the same packing as the global minimum in the AMBER force field. The distribution of structures is given in Figure 1d. A larger number of low-energy structures was found using this force field than when using the others. Clustering of those DREIDING structures with our clustering method[27] using tolerances of $\epsilon_N = 0.3$ and $\epsilon_A = 0.3$ Å did not remove any equivalent structures, which shows that this difference was not caused by a more crude clustering of the UPACK results. Apparently, the DREIDING force field had more relatively low-energy minima, an observation for which we have no good explanation.

COMPARISON POLYMORPH PREDICTOR AND UPACK

A comparison of the sampling of the UPACK and Polymorph Predictor programs was performed. The two search strategies were used with different force fields and direct clustering of the two sets of structures is nonfeasible: corresponding structures would differ too much, as a result of the inherent force field differences. To be able to cluster the two sets, the Polymorph Predictor structures were transferred to the UPACK program, minimized in the GROMOS force field, and clustered. The resulting lists were clustered with the lists that were produced with the UPACK program itself, using clustering tolerances of $\epsilon_N = 0.3$ and $\epsilon_A = 0.3$ Å. The clustering was performed in two ways, to test whether structures found by the Polymorph Predictor were not present in the UPACK list, and vice versa. The distribution of those missed structures is given in Figure 2. It can be seen that the sampling by UPACK was rather complete, as only a few low-energy structures were missed by UPACK. This gave confidence that the proper sampling grid and clustering tolerances were used. Conclusions considering the sampling of the Polymorph Predictor should be drawn with care, because the structures created by this program were completely minimized and clustered using the DREIDING force field before being transferred to UPACK and minimized in GROMOS. Thus, comparison of sampling assumes that each GROMOS force field minimum can be reached from its related minimum in the DREIDING force field. Often this will be the case, but there is no reason why this should always be so. On this basis the results shown in Figure 2b seem to be satisfactory, as only a modest number of low-energy structures was missed. A fair test would be to convert the UPACK structures to the Polymorph Predictor and minimize them in the DREIDING force field, but there seems to be no easy way to do this.

SYMMETRY CONSIDERATIONS

As an acetic acid molecule has internal symmetry (mirror plane), a crystal structure can gain symmetry during minimization: the final crystal structure can be described in another space group with the molecule on a special position. Indeed, this happened in a number of cases. An example is the structure with the lowest lattice energy in AMBER, which was also found among the relatively stable structures in the other force fields. This is a *Pnma* structure, with the acetic acid molecule on a crystallographic mirror plane (see Fig. 3). As an indication of the thoroughness of the search procedure it is worth mentioning that this structure was found in all space groups considered that are crystallographic subgroups of *Pnma* ($P2_1/c$, $P2_12_12_1$, and $Pna2_1$).

Equivalent structures of higher symmetry are not automatically clustered when they are found in different space groups. Within one space-group list it may also happen that two settings of such a structure are not clustered; that is, the transformation needed for the clustering may be forbidden in the assumed space group of lower symmetry. Within each space group these redundancies are removed by applying the distance-list-based clustering.

All structures less than 10 kJ/mol higher in energy than the global minimum in the GROMOS force field were expanded to $P1$ unit cells and subsequently minimized using a cut-off radius of 12 Å. As no rigid-body minimization was imple-

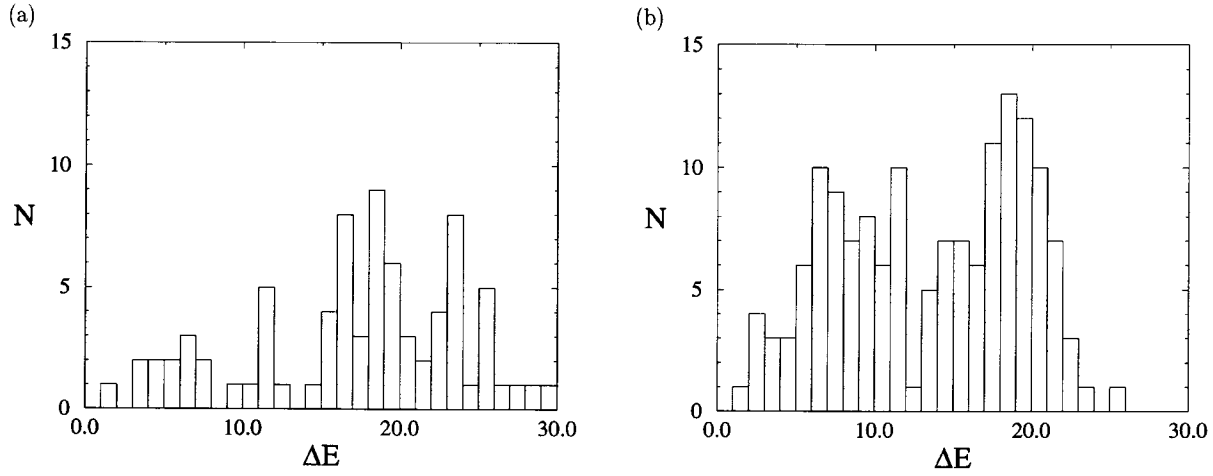

FIGURE 2. Similarity of UPACK and Polymorph Predictor sampling (ΔE is the energy difference with the global minimum (kJ/mol), N is the number of structures). (a) Polymorph Predictor structures missed by UPACK. (b) UPACK structures missed by Polymorph Predictor.

mented for more than one independent molecule, this was done with full geometry relaxation. During such a minimization the structures can lose some of their symmetry. The interesting result is that large differences between space groups occurred. Some space groups, like $P2_1$ and $P2_12_12_1$, maintained their original symmetry in all cases. However, especially in *Pnma*, *Pbca* and $C2/c$, large numbers of structures lost some of their symmetry. Those structures ended up in other space groups having two or more independent molecules. However, quite a number of structures ended up again having one independent molecule in a different space group. For example, the structure with the lowest energy to come from the list of space group *Pnma* had one independent molecule, with the space group changed to $P2_1/c$—this was one of the most stable structures already found in that space group. In the space group *Pnma* no good packing appeared to be possible with a molecule in a general position, as only relatively high-energy structures were found and, after minimization in $P1$, energies dropped considerably. The entries in the lists of both *Pbca* and $C2/c$ already started at reasonably low energy, but also they quite often lost their original symmetry, again sometimes ending up in $P2_1/c$ with one independent molecule.

All the resulting lists for different space groups were analyzed for their final symmetry using PLATON.[30] Structures having lost their original symmetry, but ending up with one independent molecule, were assigned to their proper space group. Equivalent structures in special positions were only kept in one arbitrarily chosen space group rather than classified in their correct space group of higher symmetry. Structures with $Z' > 1$ were separated from the $Z' = 1$ structures. The sorted lists for all the space groups were clustered together ($\rho_c = 7$ Å, $\eta = 0.2$ Å). The numbers of structures remaining and the number of structures with more than one independent molecule that were removed from each space-group list are given in Table IV. The effect of this symmetry check on the distribution of the structures is given in Figure 4. The decrease in the number of structures was quite significant.

As the DMAREL program is not able to constrain the space-group symmetry the results of the DMAREL minimizations were analyzed for their final symmetry, based on PLATON[30] calculations and clustered together ($\rho_c = 7$ Å, $\eta = 0.2$ Å). The numbers are given in Table IV. Again, many of the structures that started in space group *Pnma* transformed to other symmetries. The minimizations resulted in structures in various space groups (e.g., $P2_1/c$, $P\bar{1}$), some with more than one molecule in the asymmetric unit. However, the structure with the lowest energy produced from the *Pnma* list still had one independent molecule, but the space group changed to *Pccn*. As this space group was not part of the search procedure, this structure had not been found previously. Other space groups in which a large number of structures lost some of their symmetry were again $C2/c$ and *Pbca*. The resulting structures had one, two, or four independent molecules. The two lowest-energy structures to come from the DMAREL minimizations in $C2/c$

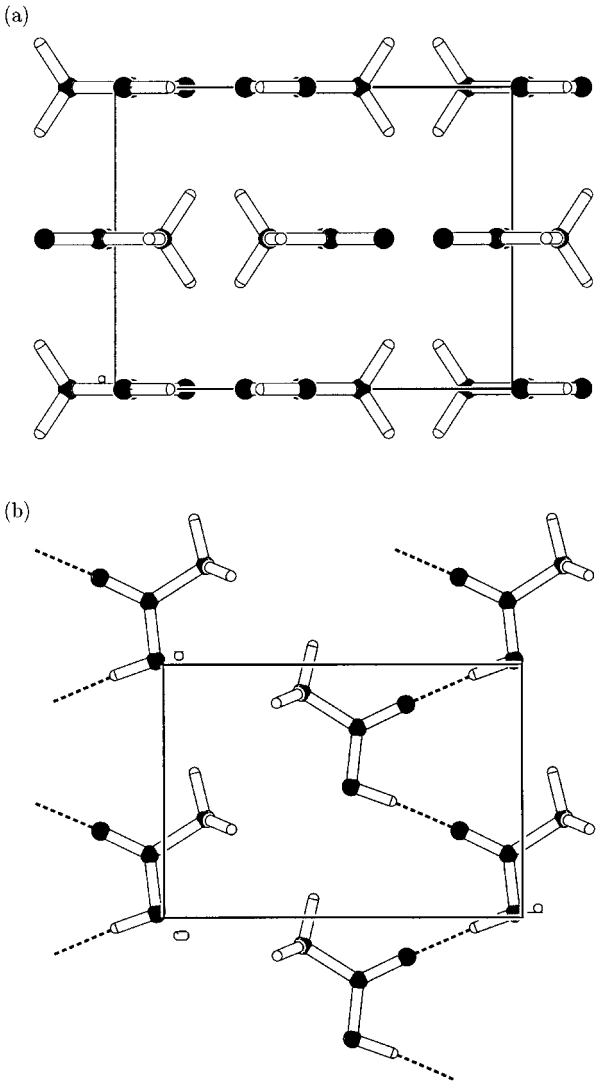

FIGURE 3. *Pnma* ($Z' = 1/2$) structure, found by minimizations with $Z' = 1$ in the space groups $Pna2_1$, $P2_12_12_1$, and $P2_1/c$. This structure had the lowest lattice energy in the AMBER force field. (a) View along the *a*-axis. (b) View along the *c*-axis.

were again fine examples of structures that ended up with $Z' = 1$ in another space group. The first one changed to space group $P2_1/c$; it was the structure corresponding to the global minimum in the lattice energy and was already found in that space group. This structure was also found starting from a structure in space group *Pbca*. The second one was the third overall ranked structure and the most stable structure found in the space group $P\bar{1}$.

In both the DMA and the GROMOS force field, no structure with more than one independent molecule had a lower lattice energy than the overall best structure with one independent molecule, but some of those structures had an energy that was only a few kilojoules per mole higher. So, it is clear that not all of those structures can be discarded on grounds of their lattice energy. However, as we limited our study to one independent molecule in the first place, we will not further discuss structures with more than one independent molecule.

MOLECULAR DYNAMICS RESULTS

The structures to come from the $P1$ minimizations (N_2 in Table IV) were subjected to a MD shake-up in the GROMOS force field. As previously discussed, the space-group symmetry was retained in those simulations. In this way, transitions between structures can take place that would be very rare in a more realistic simulation of a larger set of independent molecules. Therefore, we cannot expect quantitative results, but we did obtain some information about the thermal stability of our hypothetical structures. Two subsequent 60-ps MD runs were performed, each followed by energy minimization and clustering. The numbers of structures per space group that were left after clustering are given in Table V. The distribution over the lattice-energy differences is given in Figure 4. The MD shake-up has done what it was meant to do: less stable structures interconverting into (already found) more stable ones. The resulting lists show fewer structures and the average lattice energy is lower.

PACKING ANALYSIS

The different packing modes of carboxylic acids have been subjected to a thorough analysis by Leiserowitz.[43] Different possible catemer motifs have been distinguished, which are shown in Figure 5. The lowest GROMOS structures are catemers and both the A and B catemer motif were found. No preference for the (experimental) A motif was found. The C type was not found: acetic acid lacks the long aliphatic chains needed to stabilize such a motif.[43] The structure corresponding to the global minimum in the GROMOS force field was of the A type (Fig. 6). Dimeric structures were also found, some of which were less than 1 kJ/mol higher in energy than the global minimum. The dimers can be packed in many different ways,

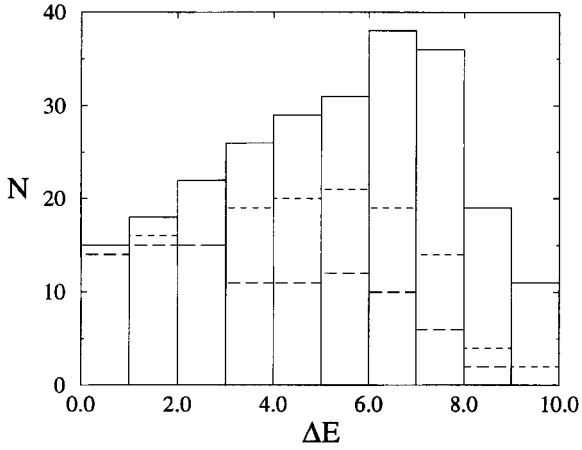

FIGURE 4. Effect of *P*1 minimization and MD on distribution of predicted structures (ΔE is the energy difference with the global minimum (kJ/mol), N is the number of structures). Solid lines: starting set; short dashed lines: after *P*1 minimization; long dashed lines: after MD.

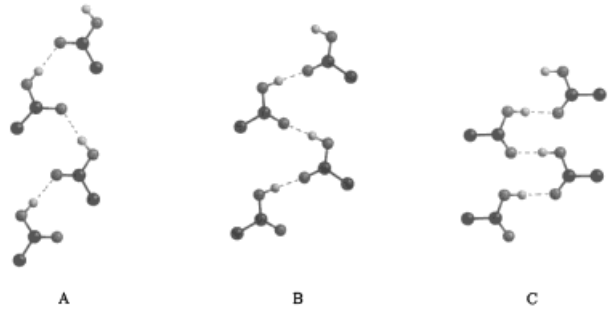

FIGURE 5. Catemer motifs as described by Leiserowitz.[43] (A) A C—O ··· O angle of ~ 120° and a translation axis of ~ 7 Å. (B) A collinear O—H ··· O bond (C—O ··· O angle of ~ 180°) and a translation axis of ~ 6.3 Å. (C) A C—O ··· O angle of ~ 120° and a translation axis of ~ 5 Å.

TABLE IV.
Results of Release of Space-Group Symmetry Constraints and Results of DMAREL Calculations.[a]

	$P\bar{1}$	$P2_1$	$P2_1/c$	$C2/c$	$P2_12_12_1$	$Pna2_1$	*Pbca*	*Pnma*
GROMOS								
N_1	4	6	77	103	16	25	31	12
N_2 (7, 0.2)	4	6	57	36	15	15	13	0
N_3	0	0	4	30	0	4	10	0
DMA								
N_1	7	13	96	150	20	40	40	21
N_2 (7, 0.2)	6	8	53	69	15	26	21	9
N_3	0	0	2	32	0	0	12	9

N_1 = number of starting structures used in minimization; N_2 = number of structures after minimization, clustering of all the lists together and selecting for correct space-group symmetry; N_3 = number of structures with $Z' > 1$, removed from this space-group list

[a] Values in parentheses denote clustering parameters ρ_c and η, respectively.

TABLE V.
Results of MD Shake-Up.[a]

	$P\bar{1}$	$P2_1$	$P2_1/c$	$C2/c$	$P2_12_12_1$	$Pna2_1$	*Pbca*	*Pnma*
N_1	4	6	57	36	15	15	13	0
N_2 (0.3, 0.3)	3	5	41	28	11	13	10	0
N_3 (0.3, 0.3)	2	5	34	26	10	13	7	0

N_1 = Number of structures at start MD shake-up; N_2 = number of structures after 60-ps MD run and clustering; N_3 = number of structures after 60-ps MD run and clustering, starting from the N_2 structures.

[a] Values in parentheses denote clustering tolerances ϵ_N and ϵ_A, respectively.[27]

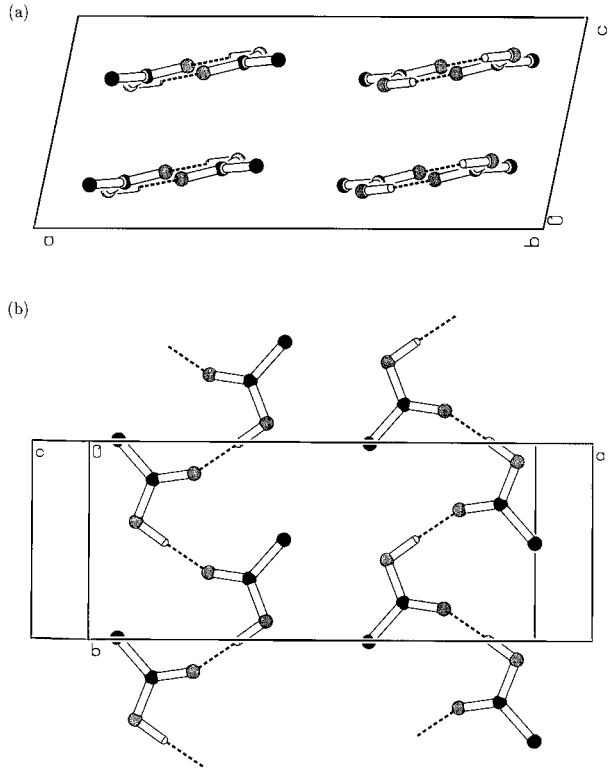

FIGURE 6. Structure that had the lowest lattice energy in the GROMOS force field ($A2/a$). (a) View along the b-axis. (b) View on the a–b plane.

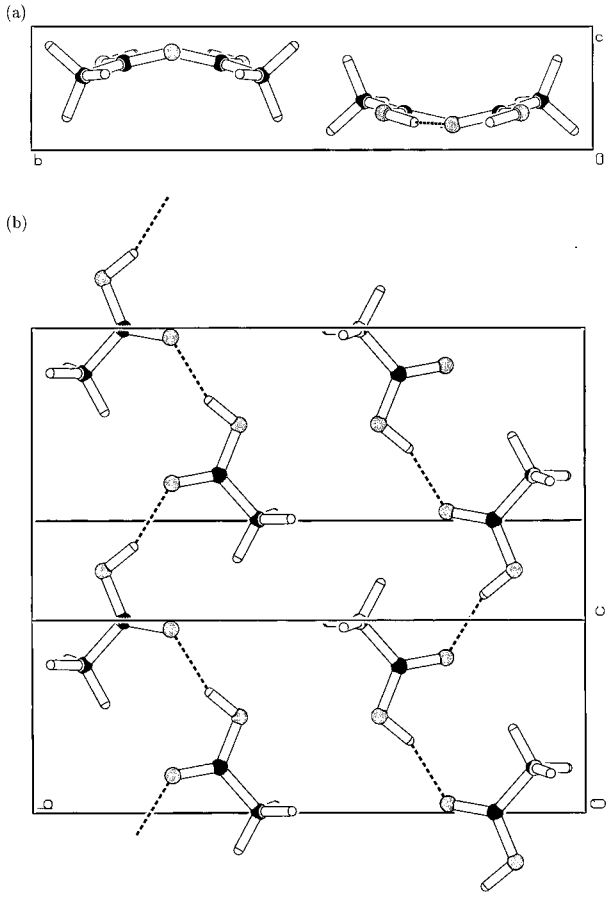

FIGURE 7. Structure that had the lowest lattice energy in the DMA force field ($P2_1/a$). (a) View along the a-axis. (b) View on the a–b plane.

being more or less recognizable as the stacking motifs given by Leiserowitz.[43] Using the AMBER force field the same kind of packing arrangements were present among the low-energy structures; the ''best'' dimer was in this case 4 kJ/mol higher in energy than the global minimum, which was a catemer of the B type (Fig. 3).

Upon DMAREL minimization the individual structures changed considerably, but overall the packing motifs observed with the DMA force field were similar to the results with the other two force fields. There was no obvious preference for the A catemer motif, although the global minimum featured this structural motif. The difference between this structure and the experimental crystal structure was in the orientation of the catemer chains with respect to each other (see Figs. 7 and 8). Dimer structures were again among the most stable ones, with the highest ranked dimer being second (less than 0.5 kJ/mol higher in energy). In the set of structures minimized in the DMA(OH) potential, the global minimum structure was a dimer. The structure ranked first in the DMA potential was now second in energy, with an energy difference of 0.3 kJ/mol. So, the relative order of structures was shifted somewhat by this alteration of potential, but still all packing motifs were present among the relatively stable structures.

C—H $\cdots$ O interactions have been considered to be a packing determining interaction.[43] Using the GROMOS force field they were not present, because of its united-atom carbon type. This is one reason why one could expect this force field to be less accurate for calculations on acetic acid. Using the AMBER or the DMA force field there were large numbers of crystal packings featuring those interactions. Many different types and geometries were found, including that present in the experimental structure: the C—H $\cdots$ O bond to the carbonyl oxygen of the first neighbor in the catemer chain. But, for example, in the first ranked structure in AMBER (also relatively stable in the DMA

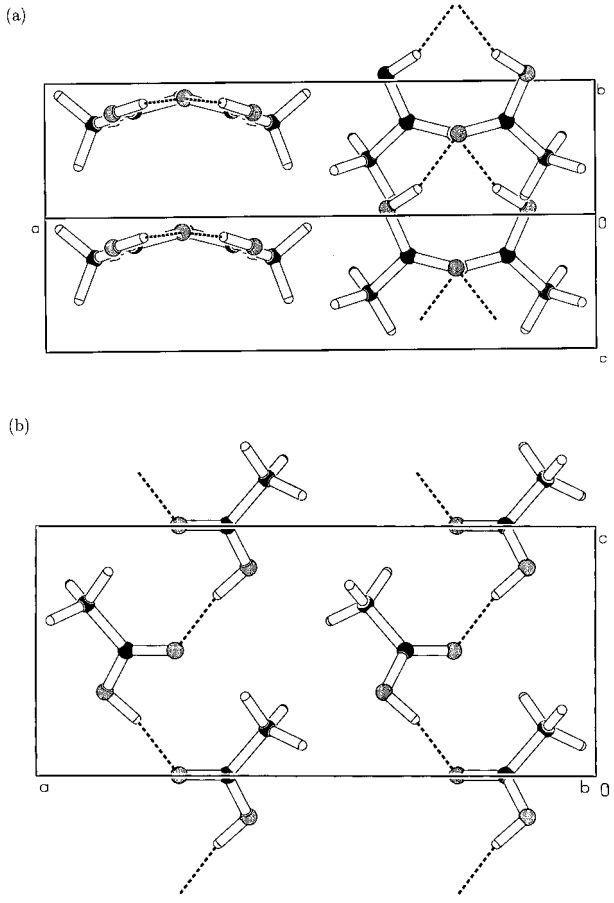

FIGURE 8. Experimental structure minimized in the DMA force field ($Pna2_1$). (a) View along the $b-c$ diagonal. (b) View along the b-axis.

force field) a different C—H ··· O interaction was present, bridging to the hydroxyl oxygen atom of the second neighbor in the catemer chain (see Fig. 3).

Discussion

The quality of a potential for crystal structure calculations can be judged by relaxing the experimental structure in the force field. When the experimental structure does not change much upon minimizing, it can be hoped that the potential is good enough to be used for extrapolating to the hypothetical structures. When the experimental structure is not very well represented, one can hardly expect the potential to give correct results for any hypothetical structures.[44]

Although the dispersion/repulsion parameters used with the DMA force field have not been optimized for this type of molecule, this potential is superior in maintaining the experimentally correct structure. Apparently, an accurate description of the anisotropy of the electrostatic interaction is very important in predicting the relative orientation of the molecules in the crystal structure. Because the DMA model gives quite an accurate description of the electrostatic potential around the *ab initio* charge distribution from which it was derived, we can have more confidence in the relative electrostatic contributions to the lattice energy, which are the larger part (typically 70%) of the total for acetic acid. Nevertheless, the overall conclusions that there are large numbers of energetical feasible structures for acetic acid and that the qualitative packing types are all stable, are essentially the same for all force fields. This is also true for the DMA(OH) potential, which gives a very satisfactory prediction of the cell dimensions and hydrogen bond geometry, within the limitations of a static model. Thus, energy minimization does not pinpoint the correct crystal structure for acetic acid.

Even the more general, older question of why acetic acid does not crystallize as a dimer cannot be answered, as very reasonable dimeric structures can be proposed. Attractive C—H ··· O interactions have been indicated to play a role in stabilizing packing arrangements of carboxylic acids.[43] Our present study has shown that a great number of structures can be proposed that incorporate such interactions, so this still controversial phenomenon cannot explain the shape of the experimental structure. Even if one would have correctly predicted both the hydrogen bonds and the C—H ··· O bonds there are still several crystal structures possible (e.g., the structure with the lowest lattice energy in the DMA force field). The experimental crystal structure cannot be explained by a single type of interaction but results from an interplay between all the different interactions, weak and strong.

Considering the instability of quite a large number of structures in the space groups $C2/c$, *Pbca*, and *Pnma*, the practice of crystal search and minimization under space-group symmetry can be questioned. It is very useful to facilitate the search procedure, as one might hardly expect a search in $P1$ with $Z' = 8$ to easily find, for example, the $C2/c$ symmetry, even though Williams[15] recently reported locating the experimental crystal struc-

ture of benzene starting from four independent molecules. One can conclude that it is necessary to subject a candidate crystal structure to a minimization in $P1$, to test whether it is stable. It is seen that, even when the energy of the starting structure is relatively low, it does not automatically mean that the structure is stable upon removal of the space-group symmetry constraints.

The short MD shake-up shows that even if the generated structures are all distinct energy minima, they can in fact be equivalent. One could try to get rid of those structures by performing a more crude clustering. However, this has the disadvantage that it is becoming increasingly uncertain whether or not two truly distinct structures may be assumed to be identical. In fact, one would like to know whether two structures are the same in the sense that they lie on the same MD trajectory. To properly determine such information, one should perform MD simulations of systems consisting of a large number of unit cells. This will be very costly in terms of computer time and there will be problems in such an analysis due to loss of symmetry. Nonetheless, we are planning to investigate further such use of molecular dynamics. As an additional advantage, some temperature effects will begin to become apparent.

A different question is what to think of all the structures with $Z' > 1$, as their energy can be rather low. Perhaps the best thing to do is to recall the statistics of the Cambridge Structural Database, which show us that only 10%[45] of all crystal structures have more than one independent molecule. Based on this percentage one can more easily concentrate on the $Z' = 1$ structures, although this number may be biased by a historical preference of crystallographers for $Z' = 1$ structures.

Careful minimization, testing for symmetry stability, and some MD shaking can reduce the number of feasible structures, but still a large number of possibilities remains. The question is raised as to what additional information should be considered to arrive at a complete *ab initio* crystal structure prediction. Of course, better force fields will result in more reliable predictions. In a molecular orbital study on the crystal structure of acetic acid, Turi and Dannenberg[46] have shown that cooperative effects play an important role. Both in the infinite hydrogen-bonding chain and in the further packing of the crystal nonadditive contributions to the lattice energy were significant and, hence, the addition of polarization terms may alter the relative energies of candidate structures. We are presently considering incorporating polarization into the model, but it is unclear how to combine that with the distributed multipole model. Adaptation of a DMA model to flexible molecules also needs to be developed.

Although the most accurate description of the intermolecular interactions is needed for crystal structure prediction it seems unlikely that a definitively accurate intermolecular potential would give the experimental crystal structure a significantly lower lattice energy than any other structure. It appears likely that, for many molecules, there is a range of energetically possible structures, and that the many other factors that affect the kinetics of crystal growth are responsible for selecting those that will be found experimentally. It will be necessary to look at such effects before one can hope to arrive at a proper prediction.

Note

After completion of this work our attention was drawn to an investigation by Payne et al.,[47] who performed crystal packing predictions for acetic acid using the Polymorph Predictor. They concluded that a generic force field like DREIDING "does not provide a good description of the hydrogen bond in molecules containing the carboxylic acid group," and, as a consequence, "it is not currently possible to have confidence in the relative stability of the predicted structure." An interesting result of that work was a proposal for the structure of the high-pressure form of acetic acid for which only powder-diffraction data are available.[48] We compared calculated powder-diffraction diagrams for low-energy structures in the DMA force field with the experimental diagram and came up with the same structure. The high-pressure form of acetic acid corresponded to the lowest-energy structure in the DMA force field (Fig. 7: $P2_1/c$, $a = 5.71$ Å, $b = 13.85$ Å, $c = 3.94$ Å, $\beta = 97.2°$). This structure gave a well-matched powder diagram, and could serve as a starting point for Rietveld refinement.

Supplementary Material

Cell parameters and fractional coordinates of the candidate crystal structures, denoted by N_3 in Table V, are included in the Supplementary Material.

Acknowledgments

We would thank Dr. A. L. Spek for assistance with and continuous improvement of the PLATON program, Mr. S. X. M. Boerrigter for preparing the first lists of structures in this prediction, Mr. D. S. Coombes for his assistance with the program DMAREL, and Dr. R. S. Payne for providing us with his manuscript prior to its publication.

References

1. A. Gavezzotti, *J. Am. Chem. Soc.*, **113**, 4622 (1991).
2. H. R. Karfunkel and R. J. Gdanitz, *J. Comput. Chem.*, **13**, 1171 (1992).
3. H. R. Karfunkel and F. J. J. Leusen, *Speedup*, **6**, 43 (1992).
4. H. R. Karfunkel, F. J. J. Leusen, and R. J. Gdanitz, *J. Comput.-Aid. Mat. Des.*, **1**, 177 (1993).
5. J. R. Holden, Z. Du, and H. L. Ammon, *J. Comput. Chem.*, **14**, 422 (1993).
6. B. P. van Eijck, W. T. M. Mooij, and J. Kroon, *Acta Cryst.*, **B51**, 99 (1995).
7. T. Shoda, K. Yamahara, K. Okazaki, and D. E. Williams, *J. Mol. Struct. (Theochem)*, **333**, 267 (1995).
8. T. Shoda and D. E. Williams, *J. Mol. Struct. (Theochem)*, **357**, 1 (1995).
9. K. D. Gibson and H. A. Scheraga, *J. Phys. Chem.*, **99**, 3765 (1995).
10. N. Tajima, T. Tanaka, T. Arikawa, T. Sakurai, S. Teramae, and T. Hirano, *Bull. Chem. Soc. Jpn.*, **68**, 519 (1995).
11. T. Arikawa, N. Tajima, S. Tsuzuki, K. Tanabe, and T. Hirano, *J. Mol. Struct. (Theochem)*, **339**, 115 (1995).
12. A. Gavezzotti, *Acta Cryst.*, **B52**, 201 (1996)
13. M. U. Schmidt and U. Englert, *J. Chem. Soc. (Dalton Trans.)*, 2077 (1996).
14. A. M. Chaka, R. Zaniewski, W. Youngs, C. Tessier, and G. Klopman, *Acta Cryst.*, **B52**, 165 (1996).
15. D. E. Williams, *Acta Cryst.*, **A52**, 326 (1996).
16. R. J. Wavak, K. D. Gibson, A. Liwo, and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA*, **93**, 1743 (1996).
17. A. V. Dzyabchenko, T. S. Pivina, and E. A. Arnautova, *J. Mol. Struct.*, **378**, 67 (1996).
18. R. J. Gdanitz, In *Theoretical Aspects and Computer Modeling of the Molecular Solid State*, A. Gavezzotti, Ed., John Wiley & Sons, New York, 1997, p. 195.
19. P. G. Jönsson, *Acta Cryst.*, **B27**, 893 (1971).
20. J. L. Derissen and P. H. Smit, *Acta Cryst.*, **A33**, 230 (1977).
21. R. J. Gdanitz, *Chem. Phys. Lett.*, **190**, 391 (1992).
22. H. R. Karfunkel, Z. J. Wu, A. Burkhard, G. Rihs, D. Sinnreich, H. M. Buerger, and J. Stanek, *Acta Cryst.*, **B52**, 555 (1996).
23. F. J. J. Leusen, *J. Cryst. Growth*, **166**, 900 (1996).
24. W. F. van Gunsteren and H. J. C. Berendsen, *GROMOS, Groningen Molecular Simulation Package*, University of Groningen, 1987.
25. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).
26. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, the Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 1989, p. 301.
27. B. P. van Eijck and J. Kroon, *J. Comput. Chem.*, **18**, 1036 (1997).
28. B. P. van Eijck and J. Kroon, *J. Phys. Chem.*, **B101**, 1096 (1997).
29. M. W. Deem, J. M. Newsam, and S. K. Sinha, *J. Phys. Chem.*, **94**, 8356 (1990).
30. A. L. Spek, *Acta Cryst.*, **A46**, C34 (1990).
31. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, *J. Am. Chem. Soc.*, **114**, 10024 (1992).
32. D. J. Tannor, B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, M. Ringnalda, W. A. Goddard III, and B. Honig, *J. Am. Chem. Soc.*, **116**, 11875 (1994).
33. A. J. Stone and M. Alderton, *Mol. Phys.*, **56**, 1047 (1985).
34. P. C. Haraiharan and J. A. Pople, *Theor. Chim. Acta*, **28**, 213 (1973).
35. R. D. Amos, I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, N. Koga, K. E. Laidig, P. E. Maslen, C. W. Murray, J. E. Rice, J. Sanz, E. D. Simandiras, A. J. Stone, and M. D. Sul, *CADPAC6: The Cambridge Analytic Derivatives Package, Issue 6*, Cambridge University, 1995.
36. D. S. Coombes, S. L. Price, D. J. Willock, and M. Leslie, *J. Phys. Chem.*, **100**, 7352 (1996).
37. D. J. Willock, S. L. Price, M. Leslie, and C. R. A. Catlow, *J. Comput. Chem.*, **16**, 628 (1995).
38. M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4907 (1977).
39. B. H. Besler, K. M. Merz, and P. A. Kollman, *J. Comput. Chem.*, **11**, 431 (1990)
40. C. H. D. C. van Ginkel, G. H. M. Calis, C. W. M. Timmermans, C. G. de Kruif, and H. A. J. Oonk, *J. Chem. Thermodyn.*, **10**, 1083 (1978).
41. A. Gavezzotti and G. Filippini, In *Theoretical Aspects and Computer Modeling of the Molecular Solid State*, A. Gavezzotti, Ed., John Wiley & Sons, New York, 1997, p. 97.
42. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, *GAUSSIAN-94, Rev. B.2*, Gaussian Inc., Pittsburgh, PA, 1995.
43. L. Leiserowitz, *Acta Cryst.*, **B32**, 775 (1976).
44. S. L. Price and K. S. Wibley, *J. Phys. Chem.*, **A101**, 2198 (1997).
45. N. Padmaja, S. Ramakumar, and M. A. Viswamithra, *Acta Cryst.*, **A46**, 725 (1990).
46. L. Turi and J. J. Dannenberg, *J. Am. Chem. Soc.*, **116**, 8714 (1994).
47. R. S. Payne, R. J. Roberts, R. C. Rowe, and R. Docherty, *J. Comput. Chem.* (in press).
48. J. E. Bertie and R. W. Wilton, *J. Chem. Phys.*, **75**, 1639 (1981).